Holistic Theoretical Model For Optimal Multiple Linear And Multiple Non Linear Regression Analysis

RAMESH CHANDRA BAGADI

ISBN: 9781692500740

DEDICATION

This text is dedicated to the all compassionate *Creator* of the *Universe*.

CONTENTS

ACKNOWLEDGMENTS

The author would like to express his deepest gratitude to all the members of his Loving Family, especially his ever loving son Rohith Rishi Bagadi, Respectable Teachers, En-Dear-Able Friends, Inspiring Social Figures, Highly Esteemed Professors, Reverence Deserving Deities that have deeply contributed in the formation of the necessary scientific temperament and the social and personal outlook of the author that has resulted in the conception, preparation and authoring of this research manuscript note. The author pays his sincere tribute to all those dedicated and sincere folk of academia, industry and elsewhere who have sacrificed a lot of their structured leisure time and have painstakingly authored treatises on Science, Engineering, Mathematics, Art and Philosophy covering all the developments from time immemorial until then, in their supreme works. It is standing on such treasure of foundation of knowledge, aided with an iota of personal god-gifted creativity that the author bases his foray of wild excursions into the understanding of natural phenomenon and forms new premises and scientifically surmises plausible laws. The author strongly reiterates his sense of gratitude and infinite indebtedness to all such '*Philosophical Statesmen*' that are evergreen personal librarians of Science, Art, Mathematics and Philosophy.

1 INTRODUCTION

The Theory Of Linear and Non Linear Regression Analysis involves modeling of data points (of the Independent Variables and the Dependent Variable) for finding any Trends, both of Linear and Non Linear kind in them. Some of these concepts are detailed in the works mentioned in the References section.

In this research manuscript, the author has advented the comprehensive Holistic Theoretical Model For Optimal Multiple Linear and Multiple Non Linear Regression Analysis. Also, an Exhaustive Error Modeling Scheme is detailed to perfect the advented Model.

2 OPTIMAL MULTIPLE LINEAR REGRESSION

Firstly, we consider Linear Regression Analysis with one independent variable. Let the independent variable be denoted by x_i and the dependent variable be denoted by y_i. The number of data points (x_i, y_i) considered are n in number. A linear relationship between x_i and y_i for any pair say (x_k, y_k) with $1 \le k \le n$ can be written as

$y_k = mx_k + c_k$ Equation 1

Therefore, we can generalize the above for all n pairs of pints as

$y_i = mx_i + c_i$ Equation 2

where $i = 1\ to\ n$ and m represents the Slope of the Straight Line and c_i represents the y ordinate.

Now, we can apply the Summation Operator on the above equation (over all n points) giving us

$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} mx_i + nc + \sum \varepsilon_i$ Equation 3a

where c is a constant.

It can be noted that in Equation 3a, we have assumed that

$nc = \sum_{i=1}^{n} c_i + \sum_{i=1}^{n} \varepsilon_i$ Equation 3b

where the ε_i are the errors in the y- ordinate value aforementioned linear approximation, when the y intercept value is supposed to be c for all the data points (x_i, y_i) when we linearly relate them with a Straight Line with Slope m.

We now divide the entire of Equation 3a by n giving us

$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} mx_i + c + \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i$ Equation 4

which can be written as

$\bar{y} = m\bar{x} + c + \bar{\varepsilon}$ Equation 5

where $\bar{y}$ and $\bar{x}$ can be computed as they are simple Arithmetic Mean values.

If we ignore the $\bar{\varepsilon}$ which is the Arithmetic Mean of the aforementioned errors, for our analysis, we can note that

$\bar{y} = m\bar{x} + c$ Equation 6

Could best represent the Average Linear Relationship between the data points x_i and y_i as the $\bar{x}$ and $\bar{y}$ are Centroids of the x_i and y_i respectively. Therefore, we are left with the task of evaluating m and c for such an aforementioned Linear Relationship.

In order to solve for m and c, we can note that we can simply use appropriate Scalar Variable Multipliers on Equation 2 and apply the Summation Operator through all the data points to generate another equation. That is, if we multiply the entire Equation 2 by x_i, and applying the Summation Operator over all n points, we have

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} m x_i^2 + c\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \varepsilon_i x_i \quad \text{Equation 7}$$

which upon division by n becomes

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i = \frac{1}{n}\sum_{i=1}^{n} m x_i^2 + \frac{c}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i x_i \quad \text{Equation 8}$$

Neglecting the last term $\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i x_i$ for the convenience of our analysis as already discussed before, we have

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i = \frac{1}{n}\sum_{i=1}^{n} m x_i^2 + \frac{c}{n}\sum_{i=1}^{n} x_i \quad \text{Equation 9}$$

which can be written as

$\overline{xy} = m(\bar{x})^2 + c\bar{x}$ Equation 10

Now, Equation 6 and Equation 10 can be solved using the theory of Simultaneous Linear Equations for the values of m and c.

However, we can note that there are infinite number of ways in which a Scalar Variable Multiplier can be producted to Equation 2 and apply the Summation Operator on all the n data points to generate another Equation.

However, in this research, the author limits the research to only certain range of such Multipliers so as to design the Optimal outcome of the m and c values. We can note that we can also multiply Equation 2 entirely by y_i, the Dependent variable to generate another equation akin to the kind of Equation 10 to solve for the values of m and c using Equation 6. The restrictions for such Scalar Variable Multipliers are detailed as follows:

Restrictions:

1. We have to use all the Variables as such Scalar Variable Multipliers at least once.
2. The highest degree of any of the terms on both sides of equation gotten by such aforementioned Scalar Variable Multipliers is equal to the sum of the number of Independent Variables and the number of Dependent Variables.

For example, if we have one Dependent Variable y_1 and two Dependent Variables x_1 and x_2, we have relations of the kind, after performing such aforementioned Scalar Variable Multiplier product operations:

$\bar{y} = m_1\bar{x}_1 + m_2\bar{x}_2 + c$ Equation 11

$\bar{y}_1^2 = m_1\bar{x}_1\bar{y}_1 + m_2\bar{x}_2\bar{y}_1 + c\bar{y}_1$ Equation 12

$\bar{y}_1\bar{x}_1 = m_1\bar{x}_1^2 + m_2\bar{x}_2\bar{x}_1 + c\bar{x}_1$ Equation 13

$\bar{y}_1\bar{x}_2 = m_1\bar{x}_1\bar{x}_2 + m_2\bar{x}_1^2 + c\bar{x}_2$ Equation 14

Since we have the degree of any of the term on either side of the Equations 12, 13 and 14 to be only 2, and the allowed level is 3 as the sum of the number of Independent and Dependent Variables.

Therefore, we can afford to multiply each of the Scalar Multiplier Variables x_1, x_2 and y_1 on each of the Equations 12, 13 and 14 once again to give us 3+3+3=9 Equations. They are:

$\bar{y}_1^3 = m_1\bar{x}_1\bar{y}_1^2 + m_2\bar{x}_2\bar{y}_1^2 + c\bar{y}_1^2$ Equation 15

$\bar{y}_1^2\bar{x}_1 = m_1\bar{x}_1^2\bar{y}_1 + m_2\bar{x}_2\bar{x}_1\bar{y}_1 + c\bar{x}_1\bar{y}_1$ Equation 16

$\bar{y}_1^2\bar{x}_2 = m_1\bar{x}_1\bar{x}_2\bar{y}_1 + m_2\bar{x}_1^2\bar{y}_1 + c\bar{y}_1\bar{x}_2$ Equation 17

$\bar{y}_1^2 x_1 = m_1\bar{x}_1^2\bar{y}_1 + m_2\bar{x}_2\bar{y}_1\bar{x}_1 + c\bar{y}_1\bar{x}_1$ Equation 18

$\bar{y}_1\bar{x}_1^2 = m_1\bar{x}_1^3 + m_2\bar{x}_2\bar{x}_1^2 + c\bar{x}_1^2$ Equation 19

$\bar{y}_1\bar{x}_1\bar{x}_2 = m_1\bar{x}_1^2\bar{x}_2 + m_2\bar{x}_2^2\bar{x}_1 + c\bar{x}_1\bar{x}_2$ Equation 20

And

$\bar{y}_1^2 x_2 = m_1\bar{x}_1\bar{x}_2\bar{y}_1 + m_2\bar{x}_2^2\bar{y}_1 + c\bar{y}_1\bar{x}_2$ Equation 21

$\bar{x}_2\bar{y}_1\bar{x}_1 = m_1\bar{x}_1^2\bar{x}_2 + m_2\bar{x}_1\bar{x}_2^2 + c\bar{x}_1\bar{x}_2$ Equation 22

$\bar{y}_1\bar{x}_2^2 = m_1\bar{x}_2^2\bar{x}_1 + m_2\bar{x}_2^3 + c\bar{x}_2^2$ Equation 23

Therefore, we now have 3 Unknowns, namely, m_1, m_2 and c and 13 Linear Equations.

$1 \rightarrow Equation\,11$

$3 \rightarrow Equations\,12,\,13\;and\;14$

$9 \rightarrow Equatons\,15,\,16,\,17,\,18,\,19,\,20,\;21,\;22,\;23$

However, we have to note that the Best 3 Equations Combination among these 13 Equations gives us the best values of m_1, m_2 and c.

Also, we have to note that since Equation 11 is the Basic Equation from which the other 12 number of Equations are built, we ought to keep this one for our use as 1 among the 3 Equations to be used for computing the best values of m_1, m_2 and c.

Therefore, we are reduced to the situation wherein we have to try to select any 2 Equations among the Equations numbered 12 through 23 using all such Combinations of 2 Equations (there would be ${}^{12}C_2$ number of such pairs) to use each pair along with Equation 11 to evaluate ${}^{12}C_2$ number of Set of values for m_1, m_2 and c. The Set of Equations that gives us the Best Values of m_1, m_2 and c is to be considered for final reporting of the Multiple Linear Regression Line. That is, this values Set gives us

the Smallest Value of the Error, i.e., $\bar{\varepsilon}$ for the Multiple Linear Regression Line

$\bar{y} = m_1\bar{x}_1 + m_2\bar{x}_2 + c + \bar{\varepsilon}$ Equation 24

In a similar fashion, we can generalize this analysis for more than 2 Independent Variables.

3 OPTIMAL MULTIPLE NON LINEAR REGRESSION

The theory of Optimal Multiple Non-Linear Regression Analysis can be performed similarly to the Optimal Linear Regression Model by applying some changed Restrictions:

Restrictions:

1. We may not need to use all the Variables as such Scalar Variable Multipliers at least once.
2. The highest degree of any of the terms on both sides of equation gotten by such aforementioned Scalar Variable Multipliers is equal to the sum of the number of Independent Variables and the number of Dependent Variables.
3. The Independent Variables are the Variables themselves and their Squares, Cubes, etc., upto (raised to the power f). That is, $fq = r$ and the variables are detailed below:

$x_1 = x_{11}, x_2 = x_{11}^2, x_3 = x_{11}^3, \ldots\ldots, x_q = x_{11}^f,$

$x_{(q+1)} = x_{21}, x_{(q+2)} = x_{21}^2, x_{(q+3)} = x_{21}^3, \ldots\ldots, x_{(q+q)} = x_{21}^f$

$x_{(2q+1)} = x_{31}, x_{(2q+2)} = x_{31}^2, x_{(2q+3)} = x_{31}^3, \ldots\ldots, x_{(2q+q)} = x_{31}^f$

………………………………………………………………………………..

………………………………………………………………………………………

…………………………………………………………………………………………..

$x_{((f-1)q+1)} = x_1, x_{((f-1)q+2)} = x_{31}^2, x_{((f-1)q+3)} = x_{31}^3, \ldots\ldots, x_{((f-1)q+q)} = x_{f1}^f$

For Example,

Say if we have

$\bar{y}_1 = m_{11}\bar{x}_1 + m_{12}\bar{x}_1^2 + m_{21}\bar{x}_2 + m_{22}\bar{x}_2^2 + c$ Equation 25

Here, the Degree of the above Equation is 2. And, the total number of Variables are 5, namely, $\bar{y}_1, \bar{x}_1, \bar{x}_1^2, \bar{x}_2, \bar{x}_2^2$. Since, the Non-Linearity is observed as relationship of the Dependent

Variable with the Independent Variables while the Independent Variables are Integral Powers of Independent Variables themselves, we consider the Degree of Equation 25 to be 2.

Hence, we can have the resulting equation after multiplying with Scalar Variables and applying the Summation Operator (on the n number of data points) in the following fashion, wherein the net degree of any of the Equation must not exceed 5.

From the Basic Equation 25 we generate 3 more Equations as follows

Multiplying Equation 25 by

$\bar{y}_1$ gives us

$$\bar{y}_1^2 = m_{11}\bar{y}_1\bar{x}_1 + m_{12}\bar{y}_1\bar{x}_1^2 + m_{21}\bar{y}_1\bar{x}_2 + m_{22}\bar{y}_1\bar{x}_2^2 + c\bar{y}_1 \quad \text{Equation 26}$$

Similarly,

Multiplying Equation 25 by

$\bar{x}_1$ gives us

$$\bar{x}_1\bar{y}_1 = m_{11}\bar{x}_1^2 + m_{12}\bar{x}_1^3 + m_{21}\bar{x}_1\bar{x}_2 + m_{22}\bar{x}_1\bar{x}_2^2 + c\bar{x}_1 \quad \text{Equation 27}$$

Multiplying Equation 25 by

$\bar{x}_2$ gives us

$$\bar{x}_2\bar{y}_1 = m_{11}\bar{x}_2\bar{x}_1 + m_{12}\bar{x}_2\bar{x}_1^2 + m_{21}\bar{x}_2^2 + m_{22}\bar{x}_2^3 + c\bar{x}_2 \quad \text{Equation 28}$$

Multiplying Equation 25 by

$\bar{x}_1^2$ gives us

$$\bar{x}_1^2\bar{y}_1 = m_{11}\bar{x}_1^3 + m_{12}\bar{x}_1^4 + m_{21}\bar{x}_1^2\bar{x}_2 + m_{22}\bar{x}_1\bar{x}_2^2 + c\bar{x}_1^2 \quad \text{Equation 29}$$

Multiplying Equation 25 by

$\bar{x}_2^2$ gives us

$$\bar{x}_2^2\bar{y}_1 = m_{11}\bar{x}_2^2\bar{x}_1 + m_{12}\bar{x}_1^2\bar{x}_2^2 + m_{21}\bar{x}_2^3 + m_{22}\bar{x}_2^4 + c\bar{x}_2^2 \quad \text{Equation 30}$$

We can note that the Highest Degree among the Set of Equations derived is still 4 and we can go upto 5.

At this stage also, we use the best 5 Equations among the 1+5=6 Equations and solve for

$m_{11}, m_{12}, m_{21}, m_{22}, c$

such that the Error is Minimum compared to any other such

combination.
Also, noting that we can Scalar Multiply the Equations 29 and 30 of Degree 4, each with $\bar{y}_1, \bar{x}_1, \bar{x}_1^2$
giving 2x3=6 more Equations and also noting that we can Scalar Multiply the Equations 26, 27 and 28 (of Degree 3) each with $\bar{x}_1^2, \bar{x}_2^2$ giving 3x2=6 more Equations. We now have a total of

$1 \rightarrow \textit{Equation}\ 25$

$5 \rightarrow \textit{Equations}\ 26,\ 27,\ 28,\ 29, 30$

$6 \rightarrow \textit{Equatons gotten by multiplying}$ Equations 29 and 30 of Degree 4, each with

$\bar{y}_1, \bar{x}_1, \bar{x}_1^2$

$6 \rightarrow \textit{Equatons gotten by multiplying}$ Equations 26, 27 and 30 of Degree 3, each with

$\bar{x}_1^2, \bar{x}_2^2$

which are a total of 18 Equations.

When Since, we have 5 unknowns and Equation 25 is the Basic Equation that generates the others, we have to form all possible groups of 4 Equations from the rest of the 17 Equations excepting Equation 25. These will be ${}^{17}C_4$ in number. Using each of these group of 4 Equations along with Equation 25, we find the Set of values of $m_{11}, m_{12}, m_{21}, m_{22}, c$ for each group. We then report that particular Set of $m_{11}, m_{12}, m_{21}, m_{22}, c$ values which best Minimizes the Error.
Therefore, it is to be noted that Clever Algebraic Degree Balancing and Scalar Multiplier Variable selection to Multiply the Equations generated with, is necessary to sustain the Highest Possible Allowed Degree in the Equations generated from the Basic Equation to solve for Optimal Non-Linear Regression Coefficients.

4 GENERALIZED MODEL FOR OPTIMAL MULTIPLE LINEAR REGRESSION

When we have r number of Independent Variables $x_1, x_2, x_3, \ldots\ldots x_{r-1}, x_r$ and 1 Dependent Variable y_1, each belonging to $\Re^n$, a linear relationship between x_i (for $i = 1\ to\ r$) and y_j (for $j = 1$) can be written as

$$y_j = \sum_{i=1}^{r} m_i x_{ji} + c \quad \text{Equation 31}$$

where we have ignored the Error ε_i for our Analysis.

Since the number of Independent Variables are r in number and Dependent Variable is 1 in number, the Highest Degree of the Equations Generated by producting Equation 31 with Scalar Variable Mutipliers, them being, $x_1, x_2, x_3, \ldots\ldots x_{r-1}, x_r$ is $(r+1)$.

We now multiply Equation 31 by x_{js} throughout where $s = 1\ to\ r$ giving us

$$y_j x_{js} = \sum_{i,s=1}^{r} m_i x_{ji} x_{js} + c x_{js} \quad \text{Equation 32}$$

Applying the Summation Operator on Equation 32 throughout for $j = 1\ to\ n$, we get

$$\sum_{j=1}^{n} y_j x_{js} = \sum_{j=1}^{n} \sum_{i,s=1}^{r} m_i x_{ji} x_{js} + c \sum_{j=1}^{n} x_{js} \quad \text{Equation 33a}$$

Dividing the Entire Equation 32 by n we get

$$\frac{1}{n} \sum_{j=1}^{n} y_j x_{js} = \frac{1}{n} \sum_{j=1}^{n} \sum_{i,s=1}^{r} m_i x_{ji} x_{js} + \frac{c}{n} \sum_{j=1}^{n} x_{js} \quad \text{Equation 33b}$$

which can be written as

$$\sum_{j=1}^{n} \bar{y}_j \bar{x}_{js} = \sum_{j=1}^{n} \sum_{i,s=1}^{r} m_i \bar{x}_{ji} \bar{x}_{js} + c \sum_{j=1}^{n} \bar{x}_{js} \quad \text{Equation 34}$$

which are s number of Equations, them being Equations (34+1) through Equations (34+s)

Also, we multiply Equation 31 throughout by y_j giving

$$y_j^2 = \sum_{i,s=1}^{r} m_i x_{ji} x_j + c x_j \quad \text{Equation (34+s+1)}$$

Applying the Summation Operator on Equation (34+s+1) throughout for $j = 1 \; to \; n$, we get

$$\sum_{j=1}^{n} y_j^2 = \sum_{j=1}^{n}\sum_{i,s=1}^{r} m_i x_{ji} x_j + c\sum_{j=1}^{n} x_j \quad \text{Equation (34+s+2)a}$$

Dividing the above Equation throughout by n, we get

$$\frac{1}{n}\sum_{j=1}^{n} y_j^2 = \frac{1}{n}\sum_{j=1}^{n}\sum_{i,s=1}^{r} m_i x_{ji} x_j + \frac{c}{n}\sum_{j=1}^{n} x_j \quad \text{Equation (34+s+2)b}$$

which can be further written as

$$\bar{y}_j^2 = m_i \bar{x}_{ji} \bar{x}_j + \frac{c}{n}\bar{x}_j \quad \text{Equation (34+s+3)}$$

Similarly, from Equation 31, we can generate the following Set of Equations

$$y_j x_{js} = \sum_{i=1}^{r} m_i x_{ji} x_{js} + c x_{js} \quad \text{Equation (34+s+4)a}$$

Also, applying the Summation Operator throughout the points $j = 1 \; to \; n$, on Equation (34+s+4) we get

$$\sum_{j=1}^{n} y_j x_{js} = \sum_{j=1}^{n}\sum_{i=1}^{r} m_i x_{ji} x_{js} + c\sum_{j=1}^{n} x_{js} \quad \text{Equation (34+s+4)b}$$

which can be further written as

$$\bar{y}_j \bar{x}_{js} = \sum_{i=1}^{r} m_i \bar{x}_{ji} \bar{x}_{js} + c\bar{x}_{js} \quad \text{Equation (34+s+4)c}$$

Also, applying the Summation Operator through the points $j = 1 \; to \; n$ on Equation 31 gives

$$\sum_{j=1}^{n} y_j = \sum_{j=1}^{n}\sum_{i=1}^{r} m_i x_{ji} + nc \quad \text{Equation (34+s+5)a}$$

Now dividing the above Equation (34+s+5)a by n, we get

$$\bar{y}_j = \sum_{i=1}^{r} m_i \bar{x}_{ji} + c \quad \text{Equation (34+s+5)b}$$

We now consider multiplying Equation (34+s+1) throughout by $y_j^{(r+1)}$ and get the following

$$y_j^{(r+1)} = \sum_{i=1}^{r} m_i x_{ji} y_j^r + c y_j^r \quad \text{Equation (34+s+5)c}$$

Actually, we can note that Equation (34+s+5) whose highest degree is 1 can be rendered into an Equation with Highest Degree $(r+1)$ in the following fashion:

That is, multiplying the following Equation (34+s+5)b

$\bar{y}_j = \sum_{i=1}^{r} m_i \bar{x}_{ji} + c$ throughout by $y^p x_i^{(r-p)}$ where $p = 0, 1, 2, 3, \ldots\ldots, r$

That is, there are r^2 number of ways which will give r^3 number of Equations as i goes from 1 to r.

Let these be Equations (34+s+7) through Equations (34+s+7+ r^3)

Though we can note that the factors $y^p x_i^{\{(r-q)-p\}}$ also give many Equations, we will get the Error Minimization while Evaluating the Regression Coefficients $m_i's$ and c only when we choose r number of best equations among the aforementioned r^3 number of Equations and the Equation (34+s+5) to solve for the r number of m_i co-efficients and the constant c which gives the Minimum Error $\bar{\varepsilon}$. The number of ways in which r number of equations can be selected from r^3 number of equations is ${}^{r^3}C_r$.

Also, we can note that Equation (34+s+5) whose highest degree is 1 can be reduced into an Equation with Highest Degree $(r+1)$ in the following fashion as well:

That is, by multiplying Equation (34+s+5) by a factor $y^{\beta_h} \prod_{i=\{\gamma\}} x_i^{\alpha_i}$ where β_h can take values from 1 to r, where

$$\sum_{i=\{\gamma_h\}} \alpha_i = (r - \beta_h) \quad \text{Equation (34+s+5)d}$$

We can note that the Cardinality of $\{\gamma_h\}$ belongs to the Set Γ detailed as below:

$\Gamma = \{1, 2, 3, \ldots\ldots, (r-1), r\}$ and $\{\gamma_h\} \subseteq \Gamma$ with $i, \beta_h, \alpha_i \in \Gamma$, i.e., i, β_h, α_i are Single Element Sub-Sets of Γ. Also, it is to be noted that we have to consider the aforementioned Multiplying Factor for all possible subsets γ_h of Γ.

This also gives us some number of additional Equations, among which including the already generated Equations we select the Best r number of Equations to use along with Equation (34+s+5) to solve for the m_i's and c. For achieving this, we have to also find all possible sets of r number of Equations from the Complete Set of aforementioned generated Equations, excepting Equation (34+s+5). Such best Set of r number of Equations when used along with Equation (34+s+5) to solve for the m_i's and c minimizes the Error $\bar{\varepsilon}$.

5 GENERALIZED MODEL FOR OPTIMAL MULTIPLE NON LINEAR REGRESSION

When we have r number of Independent Variables and each of them is used a Polynomial of Degree d_i where i goes from $i = 1\ to\ r$, then the Non-Linear Relationship between the Independent and Dependent Variables can be written as follows:

$$y_i = \sum_{i=1}^{r} \sum_{j=1}^{\alpha_\mu} m_{ji} x_{ji}^{j} + c \quad \text{Equation (34+s+6)}$$

where $\alpha_\mu \in N$, i.e., the Set of Positive Integers starting from 1 and upto μ. That is α_i is a function from the Set $\{\mu\} \to \{N\}$, for small N. We can note that the Degree of Equation (34+s+6) is $\underset{\substack{i=1\ to\ r \\ j=1\ to\alpha_\mu}}{Max}(j)$. And the allowable Degree to which we can Modify Equation (34+s+6) by applying Scalar Variable Multipliers y_i, x_{ji}^{j}, is $\left\{1 + \sum_{\mu=1}^{r} \alpha_\mu\right\}$ which is the total number of Variables, both Independent and Dependent. This is also the number of Scalar Variable Multipliers to begin with to apply these on Equation (34+s+6). These $\left\{1 + \sum_{\mu=1}^{r} \alpha_\mu\right\}$ number of Scalar Variable Multipliers when acted on Equation (34+s+6) gives us $\left\{1 + \sum_{\mu=1}^{r} \alpha_\mu\right\}$ number of Equations in addition to Equation (34+s+6). Though the situation is that of wherein the number of unknowns are less than the number of Equations available to solve for the Regression Coefficients $m_i's$ and c, we should note that this would not give us an Optimal Value of the Regression Coefficients that Minimizes the Error $\bar{\varepsilon}$.

Therefore, we consider each of the $\left\{1 + \sum_{\mu=1}^{r} \alpha_\mu\right\}$ number of

Equations that are generated from the Basic Equation (34+s+6) and multiply each of the $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$ number of Scalar Variable Multipliers to give us now $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}^{2}$ number of Equations. Now, we again check if the Degree of these Equations is less than the number $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$ or not. If not, we repeat the process of multiplying each of the $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$ Scalar Variable Multipliers with each of the Equation generated fro the Basic Equation until now, which are $\left\{1+\sum_{i=1}^{r}\alpha_{i}\right\}+\left\{1+\sum_{i=1}^{r}\alpha_{i}\right\}^{2}$ giving us $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}+\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}^{2}+\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}^{3}$ number of Equations. We keep repeating this procedure till we get the latest Set of generated Equations all having Degree greater than $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$. Once, we achieve this, we consider all those generated equations whose Degree does not exceed $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$ and find all possible groups of $\left\{1+\sum_{\mu=1}^{r}\alpha_{\mu}\right\}$ number of generated equations from all the generated equations whose Degree does not Exceed $\left\{1+\sum_{i=1}^{r}\alpha_{i}\right\}$. We solve for the Regression Coefficients m_{ji}'s and c using each such aforementioned group of generated equations and equation (34+s+6). We report those values of m_{ji}'s and c as the Regression Coefficients for which the Error $\bar{\varepsilon}_{i}$ is Minimum.

We can also note that in addition, we can also concoct Scalar Variable Multipliers in the following fashion to generate Equations from the Basic Equation:

We consider construction of the Scalar Variable Multipliers of the form

$y^{\beta_h} \prod_{\nu=\{\delta_h\}} \prod_{i=\{\theta_h\}} \prod_{j=\{\alpha_{\mu h}\}} \left(x_{ji}^{j}\right)^{\nu}$ where

$\delta_h \in N$ with small N.

$i = \{\theta_h\} \subseteq \{1, 2, 3, \ldots\ldots, (r-1), r\}$

$j = \{\alpha_{ih}\} \subseteq N$ with small N

$\nu = \{\delta_h\} \subseteq N$ with small N

$\gamma_h \subseteq \Gamma$

and $R = \sum_{\mu=1}^{r} \alpha_\mu$

where $\left\{ \sum_{j} \sum_{\nu} (j\nu) + \beta_h + \underset{i=1\ to\ r}{Max}(\alpha_\mu) \right\} = \left\{ 1 + \sum_{\mu=1}^{r} \alpha_\mu \right\}$ Equation (34+s+7)

and $\Gamma = \{1, 2, 3, \ldots\ldots, (R-1), R\}$ and $\beta_h, \alpha_i \in \Gamma$, that is β_h, α_i are Single Element Sub-Sets of Γ. Also, it is to e noted that we have to consider evaluation of the aforementioned Multiplying Factors for all possible Sub-Sets $\{\theta_h\}$, $\{\alpha_{ih}\}$ and $\{\delta_h\}$. This also gives us some number of additional Equations among which (including the already generated Equations), we select the Best r Equations to use along with Equation (34+s+6) to solve for the m_{ji}'s and c. For achieving this, we have to also find all possible Sets of and R number of Equations from the Complete Set of aforementioned generated Equations, excepting Equation (34+s+6). Such best Set of R Equations when used along with Equation (34+s+6) to solve for m_{ji}'s and c Minimizes the Error $\bar{\varepsilon}$. (Needless to mention, we have to use all R values in order to achieve Optimization. Furthermore, and finally, we can note that Equation (34+s+6) whose highest Degree is 1 can be

rendered into an Equation with Highest Degree $(R+1)$ in the following fashion:

That is, by multiplying the Equation (34+s+6) throughout by a factor $y^{\beta_h} \prod_{\nu=\{\delta_h\}} \prod_{i=\{\theta_h\}} \prod_{j=\{\alpha_{\mu h}\}} \left(x_{ji}^{j}\right)^{\nu}$ only but satisfying the below given constraint:

$$\left\{\sum_{j}\sum_{\nu}(j\nu)+\beta_h+\underset{i=1\,to\,r}{Max}(\alpha_\mu)\right\}=\left\{1+\sum_{\mu=1}^{r}\alpha_\mu\right\}-z$$ Equation (34+s+8)

with z taking all possible Positive Integer Values for each h. We again now find the Optimal Regression Coefficients m_{ji}'s and c that best Minimizes the Error $\bar{\varepsilon}$.

6 ERROR ANALYSIS

The same Error Metrics such as in the conventional Multiple Linear Regression Analysis and Multiple Non-Linear Regression Analysis can be used for the Optimal cases as well.

Exhaustive Modeling Of Error And According Updation Of The Model

The Errors for all $\bar{\varepsilon}_i$ for all (x_i, y_i) which are n in number, for the afore-described Model are again modeled using the same Model again using the following transformations for the updation of the Model to include or explain the Error:

Case 1: Optimal Multiple Linear Regression

$m_{\psi\sigma} = \omega_{pq} m_{ij}$ with $p \neq i,\ q \neq (j,\ p = 1\, to n)$ Equation (34+s+9)

which consequently leads to the transformation relation $(m_{\psi\sigma} = \omega_{pq} m_{ij}) \rightarrow \{m_{\psi\sigma} = \omega_{pq}(m_{ij} + \kappa)\}$ Equation (34+s+10)

and also, $c = \omega_c m_{ij}$ Equation (34+s+10)

leading to the transformation relation $(c = \omega_c m_{ij}) \rightarrow \{c = \omega_c (m_{ij} + \kappa)\}$. Equation (34+s+11)

Case 2: Optimal Multiple Non Linear Regression

$m_{\psi\sigma} = \sum_{\mu=1}^{N} \{(\omega_{ij})^{\alpha_\mu} m_{ij}\}$ with $p \neq i,\ q \neq (j,\ p = 1\, to n)$ Equation (34+s+12)

which consequently leads to the transformation relation $\left\{m_{\psi\sigma} = \sum_{\mu=1}^{N} \{(\omega_{ij})^{\alpha_\mu} m_{ij}\}\right\} \rightarrow \left\{m_{\psi\sigma} = \sum_{\mu=1}^{N} \{(\omega_{ij})^{\alpha_\mu} (m_{ij} + \kappa)\}\right\}$ Equation (34+s+13)

and also, $c = \sum_{\mu=1}^{N} \left\{ \left(\omega_{ij} \right)^{\alpha_\mu} m_{ij} \right\}$ leading to the transformation relation $\left\{ c = \sum_{i=1}^{N} \left\{ \left(\omega_{ij} \right)^{\alpha_\mu} m_{ij} \right\} \right\} \to \left\{ c = \sum_{i=1}^{N} \left\{ \left(\omega_{ij} \right)^{\alpha_\mu} \left(m_{ij} + \kappa \right) \right\} \right\}$,

Equation (34+s+14)

where N is a Set Of Positive Integers with small N·

We keep repeating this procedure again and again till we achieve Error Convergence after some such Steps of Updation of Model for the case of Optimal Multiple Linear Regression Model and till we achieve Zero Error after some Steps of Updation of Model for the case of Optimal Multiple Linear Regression Model.

6 REFERENCES

Rencher, Alvin C.; Christensen, William F. (2012), "Chapter 10, Multivariate regression – Section 10.1, Introduction", Methods of Multivariate Analysis, Wiley Series in Probability and Statistics, 709 (3rd ed.), John Wiley & Sons, p. 19, ISBN 9781118391679.

Goldstein, H. (1986). "Multilevel Mixed Linear Model Analysis Using Iterative Generalized Least Squares". Biometrika. 73 (1): 43–56. doi:10.1093/biomet/73.1.43. JSTOR 2336270

Yan, Xin (2009), Linear Regression Analysis: Theory and Computing, World Scientific, pp. 1–2, ISBN 9789812834119

David A. Freedman (2009). Statistical Models: Theory and Practice. Cambridge University Press. p. 26.

ABOUT THE AUTHOR

Mr. Ramesh Chandra Bagadi is currently splitting his time as a *Technology Entrepreneur* developing solutions to latest cutting edge technologies of the futuristic kind and as an *Author Of Self Help Books.* He has received his Bachelors of Civil Engineering from Osmania University, India and a Masters each in Engineering Mechanics, Civil & Environmental Engineering, Physics from the University Of Wisconsin-Madison, USA. He is also a Registered Chartered Engineer and Fellow of The Institution Of Engineers, India.